

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/709,905A

DATE: 12/13/2001

TIME: 20:27:05

Input Set : N:\Crf3\Refhold\I709905A.raw

Output Set: N:\CRF3\12132001\I709905A.raw

ENTERED

RECEIVED

FEB 06 2002

TECH CENTER 1600/2900

1631 0420/600

1 <110> APPLICANT: Kalyanaraman Ramnarayan
 2 Edward T. Maggio
 3 P. Patrick Hess
 4 <120> TITLE OF INVENTION: Use of Computationally Derived Protein
 5 Structures of Genetic Polymorphisms in Pharmacogenomics for
 6 Drug Design and Clinical Applications
 7 <130> FILE REFERENCE: 24737-1906C
 C--> 8 <140> CURRENT APPLICATION NUMBER: US/09/709,905A
 9 <141> CURRENT FILING DATE: 2000-11-10
 10 <150> PRIOR APPLICATION NUMBER: 09/438,566
 11 <151> PRIOR FILING DATE: 1999-11-10
 12 <150> PRIOR APPLICATION NUMBER: 24737-1906B
 13 <151> PRIOR FILING DATE: 2000-11-01
 14 <160> NUMBER OF SEQ ID NOS: 118
 15 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 17 <210> SEQ ID NO: 1
 18 <211> LENGTH: 6
 19 <212> TYPE: PRT
 20 <213> ORGANISM: Artificial Sequence
 21 <220> FEATURE:
 22 <223> OTHER INFORMATION: Modified Hepatitis C Virus NS3 Protease Inhibitor
 23 Peptide
 24 <221> NAME/KEY: ACETYLATION
 25 <222> LOCATION: 1
 26 <221> NAME/KEY: MOD_RES
 27 <222> LOCATION: 2
 28 <223> OTHER INFORMATION: D-glutamic acid
 29 <221> NAME/KEY: MOD_RES
 30 <222> LOCATION: 5
 31 <223> OTHER INFORMATION: beta-cyclohexylalanine
 32 <300> PUBLICATION INFORMATION:
 33 <301> AUTHORS: Ingallinella, P., Altamura, S., Bianchi, E., Talia
 34 <302> TITLE: Potent Peptide Inhibitors Of Human Hepatitis C Vir
 35 <303> JOURNAL: Biochemistry
 36 <304> VOLUME: 37
 37 <305> ISSUE: 25
 38 <306> PAGES: 8906-8914
 39 <307> DATE: 1998-06-23
 40 <400> SEQUENCE: 1
 W--> 41 Asp Xaa Leu Ile Xaa Cys
 42 1 5
 44 <210> SEQ ID NO: 2
 45 <211> LENGTH: 6
 46 <212> TYPE: PRT
 47 <213> ORGANISM: Artificial Sequence
 48 <220> FEATURE:
 49 <223> OTHER INFORMATION: Modified Hepatitis C Virus NS3 Protease Inhibitor

RAW SEQUENCE LISTING

DATE: 12/13/2001

PATENT APPLICATION: US/09/709,905A

TIME: 20:27:05

Input Set : N:\Crf3\Refhold\I709905A.raw

Output Set: N:\CRF3\12132001\I709905A.raw

50 Peptide

51 <221> NAME/KEY: ACETYLTATION

52 <222> LOCATION: 1

53 <221> NAME/KEY: MOD_RES

54 <222> LOCATION: 5

55 <223> OTHER INFORMATION: beta-cyclohexylalanine

56 <300> PUBLICATION INFORMATION:

57 <301> AUTHORS: Ingallinella, P., Altamura, S., Bianchi, E., Talia

58 <302> TITLE: Potent Peptide Inhibitors Of Human Hepatitis C Vir

59 <303> JOURNAL: Biochemistry

60 <304> VOLUME: 37

61 <305> ISSUE: 25

62 <306> PAGES: 8906-8914

63 <307> DATE: 1998-06-23

64 <400> SEQUENCE: 2

W--> 65 Asp Glu Leu Ile Xaa Cys

66 1 5

68 <210> SEQ ID NO: 3

69 <211> LENGTH: 1045

70 <212> TYPE: DNA

71 <213> ORGANISM: Human Immunodeficiency Virus (HIV)

72 <220> FEATURE:

73 <221> NAME/KEY: CDS

74 <222> LOCATION: (1)...(297)

75 <223> OTHER INFORMATION: Protease

76 <221> NAME/KEY: CDS

77 <222> LOCATION: (298)...(1045)

78 <223> OTHER INFORMATION: Portion of Reverse Transcriptase

79 <400> SEQUENCE: 3

80 cct cag atc act ctt tgg caa cga ccc cty gtc aca ata aag ata ggg 48

W--> 81 Pro Gln Ile Thr Leu Trp Gln Arg Pro Xaa Val Thr Ile Lys Ile Gly

82 1 5 10 15

83 ggc caa cta aaa gaa gct yta tta gat aca gga gca gat gat aca gta 96

W--> 84 Gly Gln Leu Lys Glu Ala Xaa Leu Asp Thr Gly Ala Asp Asp Thr Val

85 20 25 30

86 tta gaa gaa atg agt tta cca ggg aaa tgg aaa cca aaa atg ata ggg 144

87 Leu Glu Glu Met Ser Leu Pro Gly Lys Trp Lys Pro Lys Met Ile Gly

88 35 40 45

89 gga att gga ggt ttt atc aaa gta aga cag tat gat caa ata ctc ata 192

90 Gly Ile Gly Gly Phe Ile Lys Val Arg Gln Tyr Asp Gln Ile Leu Ile

91 50 55 60

92 gaa atc tgt gga cat aaa gct ata ggc aca gta tta gta gga cct aca 240

93 Glu Ile Cys Gly His Lys Ala Ile Gly Thr Val Leu Val Gly Pro Thr

94 65 70 75 80

95 cct gtc aac ata att gga aga aat ttg ttg act cag att ggt tgc act 288

96 Pro Val Asn Ile Ile Gly Arg Asn Leu Leu Thr Gln Ile Gly Cys Thr

97 85 90 95

98 tta aat ttg ccc att agt cct att gaa act gta cca gta aaa tta aag 336

99 Leu Asn Leu Pro Ile Ser Pro Ile Glu Thr Val Pro Val Lys Leu Lys

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/709,905A

DATE: 12/13/2001

TIME: 20:27:05

Input Set : N:\Crf3\Refhold\I709905A.raw

Output Set: N:\CRF3\12132001\I709905A.raw

100	100	105	110	
101	cca gga atg gat ggc cca aaa gtt aaa caa tgg cca ttg aca gaa gaa			384
102	Pro Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu Glu			
103	115 120 125			
104	aaa ata aaa gca tta gta gaa att tgt aca gaa atg gaa aag gaa gga			432
105	Lys Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met Glu Lys Glu Gly			
106	130 135 140			
107	aaa att tca aaa att ggg cct gag aat cca tac aat act cca ata ttt			480
108	Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Ile Phe			
109	145 150 155 160			
110	gcc ata aag aaa aaa gac agt act aaa tgg aga aaa tta gta gat ttc			528
111	Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe			
112	165 170 175			
113	aga gaa ctt aat aag aga aca caa gac ttc tgg gaa gtt caa tta gga			576
114	Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly			
115	180 185 190			
116	ata cca cac ccc gca ggg tta aaa cag aaa aaa tca gta aca ata ctg			624
117	Ile Pro His Pro Ala Gly Leu Lys Gln Lys Lys Ser Val Thr Ile Leu			
118	195 200 205			
119	gat gtg ggt gat gca tat ttt tca gtt ccc tta gat gaa ggc ttc agg			672
120	Asp Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Gly Phe Arg			
121	210 215 220			
122	aag tat act gca ttt acc ata cct agt aga aat aat gag aca cca ggg			720
123	Lys Tyr Thr Ala Phe Thr Ile Pro Ser Arg Asn Asn Glu Thr Pro Gly			
124	225 230 235 240			
125	att aga tat cag tac aac gtg ctc cca cag gga tgg aaa gga tca cca			768
126	Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Ser Pro			
127	245 250 255			
128	gca ata ttt caa agt agc atg aca aga aty tta gag cct ttt aga aaa			816
129	Ala Ile Phe Gln Ser Ser Met Thr Arg Xaa Leu Glu Pro Phe Arg Lys			
130	260 265 270			
131	caa aat cca gaa ata gtt atc tat caa tac atg gat gat ttg tat gta			864
132	Gln Asn Pro Glu Ile Val Ile Tyr Gln Tyr Met Asp Asp Leu Tyr Val			
133	275 280 285			
134	gga tct gac tta gaa ata gga cag cat aga gca aaa ata gag gaa ctg			912
135	Gly Ser Asp Leu Glu Ile Gly Gln His Arg Ala Lys Ile Glu Glu Leu			
136	290 295 300			
137	aga gga cat cta tta aag tgg gga ttt acc aca cca gac aaa aaa cat			960
138	Arg Gly His Leu Leu Lys Trp Gly Phe Thr Thr Pro Asp Lys Lys His			
139	305 310 315 320			
140	cag aaa gaa cct cca ttt ctt tgg atg ggt tat gaa ctc cat cct gat			1008
141	Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro Asp			
142	325 330 335			
143	aaa tgg aca gta cag cct ata aag ttg cca gaa aaa g			1045
144	Lys Trp Thr Val Gln Pro Ile Lys Leu Pro Glu Lys			
145	340 345			
147	<210> SEQ ID NO: 4			
148	<211> LENGTH: 1046			
149	<212> TYPE: DNA			

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/709,905A

DATE: 12/13/2001

TIME: 20:27:05

Input Set : N:\Crf3\Refhold\I709905A.raw

Output Set: N:\CRF3\12132001\I709905A.raw

150 <213> ORGANISM: Human Immunodeficiency Virus (HIV)
 151 <220> FEATURE:
 152 <221> NAME/KEY: CDS
 153 <222> LOCATION: (1)...(297)
 154 <223> OTHER INFORMATION: HIV Protease
 155 <221> NAME/KEY: CDS
 156 <222> LOCATION: (298)...(1046)
 157 <223> OTHER INFORMATION: Portion of HIV Reverse Transcriptase
 158 <400> SEQUENCE: 4

159	cct cag atc act ctt tgg caa cga ccc ctt gtc aca ata aag ata gga	48
160	Pro Gln Ile Thr Leu Trp Gln Arg Pro Leu Val Thr Ile Lys Ile Gly	
161	1 5 10 15	
162	ggg cag cta aag gaa gct cta tta gat aca gga gca gat gat aca gta	96
163	Gly Gln Leu Lys Glu Ala Leu Leu Asp Thr Gly Ala Asp Asp Thr Val	
164	20 25 30	
165	gtt gaa gaa atg aat ttg cca gga aaa tgg aaa cca aaa atg ata ggg	144
166	Val Glu Glu Met Asn Leu Pro Gly Lys Trp Lys Pro Lys Met Ile Gly	
167	35 40 45	
168	gga att gga ggt ttt atc aaa gta aga cag tat gag caa ata gcc gta	192
169	Gly Ile Gly Gly Phe Ile Lys Val Arg Gln Tyr Glu Gln Ile Ala Val	
170	50 55 60	
171	gaa aty tgt gga cat aga gct atg ggt aca gta tta gta gga cct aca	240
172	Glu Xaa Cys Gly His Arg Ala Met Gly Thr Val Leu Val Gly Pro Thr	
173	65 70 75 80	
174	cct gtc aac ata att gga aga aat ctg ttg act cag att ggt tgc act	288
175	Pro Val Asn Ile Ile Gly Arg Asn Leu Leu Thr Gln Ile Gly Cys Thr	
176	85 90 95	
177	tta aat ttt ccc att agt cct att gaa act gta cca gta aaa tta aag	336
178	Leu Asn Phe Pro Ile Ser Pro Ile Glu Thr Val Pro Val Lys Leu Lys	
179	100 105 110	
180	cca gga atg gat ggc cca aaa gtt aaa caa tgg cca ttg aca gaa gaa	384
181	Pro Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu Glu	
182	115 120 125	
183	aaa ata aaa gca tta gta gaa atc tgt aca gaa ttg gaa aag gaa ggg	432
184	Lys Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Leu Glu Lys Glu Gly	
185	130 135 140	
186	aaa att tca aaa att ggg cct gaa aat cca tac aat act cca gta ttt	480
187	Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe	
188	145 150 155 160	
189	gcc ata aag aaa aag aac agt act aaa tgg aga aaa tta gta gat ttc	528
190	Ala Ile Lys Lys Lys Asn Ser Thr Lys Trp Arg Lys Leu Val Asp Phe	
191	165 170 175	
192	aga gaa ctt aat aag aga act caa gac ttc tgg gag gtt caa tta gga	576
193	Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly	
194	180 185 190	
195	ata cca cat cca gca ggg tta aaa aag aat aaa tca ata aca gta ctg	624
196	Ile Pro His Pro Ala Gly Leu Lys Lys Asn Lys Ser Ile Thr Val Leu	
197	195 200 205	
198	gat gtg ggt gat gca tat ttt tca gtt ccc tta tgt gaa gac ttc agg	672

RAW SEQUENCE LISTING

DATE: 12/13/2001

PATENT APPLICATION: US/09/709,905A

TIME: 20:27:05

Input Set : N:\Crf3\Refhold\I709905A.raw

Output Set: N:\CRF3\12132001\I709905A.raw

```

199   Asp Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Cys Glu Asp Phe Arg
200       210                      215                      220
201   aag tat act gca ttt acc ata cct agt gta aac aat gag act cca ggg      720
202   Lys Tyr Thr Ala Phe Thr Ile Pro Ser Val Asn Asn Glu Thr Pro Gly
203       225                      230                      235                      240
204   att aga tat cag tac aat gtg ctt cca cag gga tgg aaa gga ttc acc      768
205   Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Phe Thr
206       245                      250                      255
207   agc ata ttc caa tgt agc atg aca aaa atc tta gag cct ttt aga aaa      816
208   Ser Ile Phe Gln Cys Ser Met Thr Lys Ile Leu Glu Pro Phe Arg Lys
209       260                      265                      270
210   caa aat cca gag ata gtt atc tat caa tac atg gat gat ttg tat gta      864
211   Gln Asn Pro Glu Ile Val Ile Tyr Gln Tyr Met Asp Asp Leu Tyr Val
212       275                      280                      285
213   gga tct gac tta gaa ata ggg cag cat aga gca aaa ata gag gaa ctg      912
214   Gly Ser Asp Leu Glu Ile Gly Gln His Arg Ala Lys Ile Glu Glu Leu
215       290                      295                      300
216   aga caa tat ctg tgg aag tgg gga ttt tgc aca cca gaa caa aar cat      960
217   Arg Gln Tyr Leu Trp Lys Trp Gly Phe Cys Thr Pro Glu Gln Lys His
218       305                      310                      315                      320
219   cag aaa gaa cct cct ttc ctt tgg atg ggt tat gaa ctc cat ccc gat      1008
220   Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro Asp
221       325                      330                      335
222   aaa tgg aca gta caa cct ata gtg ctg cca gac aaa ga      1046
223   Lys Trp Thr Val Gln Pro Ile Val Leu Pro Asp Lys
224       340                      345
226 <210> SEQ ID NO: 5
227 <211> LENGTH: 1104
228 <212> TYPE: DNA
229 <213> ORGANISM: Human Immunodeficiency Virus (HIV)
230 <220> FEATURE:
231 <221> NAME/KEY: CDS
232 <222> LOCATION: (1)...(297)
233 <223> OTHER INFORMATION: HIV Protease
234 <221> NAME/KEY: CDS
235 <222> LOCATION: (298)...(1104)
236 <223> OTHER INFORMATION: Portion of HIV Reverse Transcriptase
237 <400> SEQUENCE: 5
238   cct cag atc act ctt tgg caa cga ccc atc gtc aca ata aag rta ggg      48
239   Pro Gln Ile Thr Leu Trp Gln Arg Pro Ile Val Thr Ile Lys Xaa Gly
240       1                      5                      10                      15
241   ggg caa cta agg gaa gct cta tta gat aca gga gca gat gat aca ata      96
242   Gly Gln Leu Arg Glu Ala Leu Leu Asp Thr Gly Ala Asp Asp Thr Ile
243       20                      25                      30
244   ata gaa gac ata act ttg cca gga aga tgg aca cca aaa atg ata ggg      144
245   Ile Glu Asp Ile Thr Leu Pro Gly Arg Trp Thr Pro Lys Met Ile Gly
246       35                      40                      45
247   gga att gga ggt ttt gtc aaa gta aga cag tat gat cag ata ccc ata      192
248   Gly Ile Gly Gly Phe Val Lys Val Arg Gln Tyr Asp Gln Ile Pro Ile

```

Use of n and/or Xaa has been detected in the Sequence Listing.
 Review the Sequence Listing to insure a corresponding
 explanation is presented in the <220> to <223> fields of
 each sequence using n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/709,905A

DATE: 12/13/2001

TIME: 20:27:06

Input Set : N:\Crf3\Refhold\I709905A.raw

Output Set: N:\CRF3\12132001\I709905A.raw

L:8 M:270 C: Current Application Number differs, Wrong Format

L:41 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1

L:65 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2

L:81 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3

L:84 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3

L:129 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3

L:172 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4

L:239 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5

L:251 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5

L:263 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5

L:411 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7

L:412 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7

L:418 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7

L:491 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8

L:494 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8

L:576 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9

L:591 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9

L:630 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9

L:636 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9

L:664 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10

L:667 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10

L:706 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10

L:712 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10

L:746 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11

L:749 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11

L:764 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11

L:800 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11

L:885 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12

L:916 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13

L:919 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13

L:952 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13

L:967 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13

L:1049 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14

L:1055 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14

L:1058 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14

L:1092 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15

L:1095 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15

L:1134 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15

L:1146 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15

L:1152 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15

L:1195 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16

L:1225 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16

L:1256 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17

L:1265 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17

L:1298 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17

L:1301 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17

L:1313 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17

L:1316 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/709,905A

DATE: 12/13/2001

TIME: 20:27:06

Input Set : N:\Crf3\Refhold\I709905A.raw

Output Set: N:\CRF3\12132001\I709905A.raw

L:1324 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17

L:1350 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18

L:1377 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18

L:2102 M:351 W: Sequence data Name/Key Feature Out-of-Range, SEQ ID#:27, CDS LOCATION:
(298)...(1116)